# **Speech**

1. ASER evaluation
2. Prosody

## **Automated Diagnostic Tool to measure ASER level**

**Objective:** Create an automated diagnostic tool to measure ASER level of children.

**Concept:** Developing a ML model that would be trained on ASER data and then be utilized as a standalone testing tool and as a part of adaptive learning solutions for basic numeracy and literacy in English and Indic languages.

**Issues addressed:** Oral evaluation is considered the most effective approach right now for basic literacy and spoken fluency. The need of trained evaluators for the same, puts a limit on the number of children that can be evaluated. Without a diagnostic evaluation it is difficult for teachers to provide learning content according to the child’s level.

**Tech Needs:** A list of technology know-how or clear projects.

|  |  |  |
| --- | --- | --- |
| **Projects** | **Know-how** | **Status** |
| An android app to collect labeled dataset to be used for testing and training the ML model | Android Development | Done |
| Audio sample filtering - remove background noise and speaker diarization to extract only child’s voice | Audio-signal processing | In-Progress |
| Develop and test ML Model to take filtered audio samples and corresponding questions/text and labels (no. of mistakes and proficiency-level) as input and predict proficiency level | Machine Learning and Deep Learning Algorithms | Not Started |
| Conceptualize and integrate the designed ML model with an Android app (Deployment) | Android Development | Not Started |

**Details:** We are moving to the era of self and group learning; an automated assessment tools would be imperative to drive the motivation and check learning outcomes.

The model will consume audio samples of a child and predict his/her basic numeracy and

literacy level. A child can be classified in any of the 5 Reading levels(Beginner, Letter level, word level, short paragraph level, story level) and 5 Numeracy levels (Beginner,

Number recognition(0-9), Number recognition(11-99), two digital subtraction

with borrowing, three digit by one digit division)

For this classification, the features used will be pronunciation correctness, speaking rate, and fluency and calculation correctness for Math problems:

● Text/pronunciation mistakes (analyzed through speech to text recognition)

● No. of pauses and length of pause (analyzed through audio signal processing)

● Calculation mistakes (for Math problems)

Growing on the initial understanding of reading fluency, we hope to create oral assessment tools for science, humanities, communication skills and more. Moreover, we can recognise children’s strengths and weaknesses and plan their courses accordingly.

**Progress:** An Android app has been designed to collect human evaluated training data. This app is being used in different villages of India(Maharashtra, Uttar Pradesh and Rajasthan) to collect audio samples of children reading out different levels of text (letter, word, paragraph and story), numbers (two digit, single digit) and solution to basic numeric problems(subtraction and division) along with the proficiency level marked by the evaluator.

## **Assessment of prosodic skills**

**Objective:** Create an automated diagnostic tool to measure prosodic skills of children’s oral reading

**Concept:** To develop test tools for the effective and comprehensive assessment of prosody of children's oral reading in English as well as Indic languages.

**Issues addressed:** Accurate perception and production of the ranging aspects of prosody are a significant component of successful social communication. The need of trained evaluators for the same, puts a limit on the number of children that can be evaluated. Without a diagnostic evaluation it is difficult for teachers to provide learning content according to the child’s need.

**Tech Needs:** A list of technology know-how or clear projects.

|  |  |  |
| --- | --- | --- |
| **Projects** | **Know-how** | **Status** |
| Project Conceptualization - which prosodic skills to look for and what should be the labels | Prosodic Skills Identification | Not Started |
| An android app to collect labeled dataset to be used for testing and training the ML model | Android Development | Not Started |
| Audio sample filtering - remove background noise and speaker diarization to extract only child’s voice | Audio-signal processing | Not Started |
| Develop and test ML Model to take filtered audio samples and corresponding labels as input and predict the prosodic skills and areas of improvement | Machine Learning and Deep Learning Algorithms | Not Started |
| Conceptualize and integrate the designed ML model with an Android app (Deployment) | Android Development | Not Started |

**Details:** Pitch, duration, and stress are the phonetic correlates of prosody. The physical correlates of these features are the speech’s fundamental frequency (F0), syllable duration, and intensity, respectively. Phonological correlates of prosody include the variations in pitch, length, and loudness that are produced in speech for conveying subtle changes in the meaning of spoken messages independent of words and grammatical order. In other words, stress and intonation are used to convey the grammatical, affective, and pragmatic functions of language.

We aim to design a model in which the functions of prosody are identified at the grammatical and affective levels of communication.

The grammatical functions of prosody include

(a) determining the boundaries of phrases, clauses, or sentences, particularly when there is ambiguity (e.g., /FRUIT, SALAD and MILK/ vs. /FRUIT-SALAD and MILK/) and

(b) differentiating between word classes when there are homonyms (e.g., a noun and verb that sound the same but are differentiated by prosody).

The affective function of prosody is the use of prosodic features such as pitch contour, pauses, and word stress to express the speaker’s emotions and attitudes. Prosodic patterns convey different emotions. Happiness, for example, is characterized by a fast speaking rate, rising pitch, high variability, and fast voice onsets, and sadness is nearly the opposite.

Such models could help teachers identify which students are making adequate progress.

Source:

<https://www.asha.org/uploadedFiles/ASHA/Publications/cicsd/2015S-Behavioral-Measures-to-Evaluate-Prosodic-Skills.pdf>

# **Vision**

1. Video evaluation
2. Communication skills
3. Custom feature extraction from video assignments

2. Object Identification/OCR

1. Word Pirate game
2. Biodiversity app

## **Evaluation of Non-Verbal Communication Skills**

**Objective:** To develop Video-based Test of Non-Verbal Communication Skills (VTCS)

**Concept:** To develop test tools for the effective and comprehensive assessment of non-verbal communication skills of children

**Issues addressed:** Non-verbal cues frequently reveal the intention of the children and reflect his/her emotional reactions to any topics/subject. Without an automated evaluation tool it is difficult for teachers to identify these non-verbal cues and guide and provide content to children according to their need.

**Tech Needs:** A list of technology know-how or clear projects.

|  |  |  |
| --- | --- | --- |
| **Projects** | **Know-how** | **Status** |
| Project Conceptualization - which non-verbal skills to look for and what should be the labels | Non-Verbal Communication Expert | Not Started |
| An android app to collect labeled dataset to be used for testing and training the ML model | Android Development | Not Started |
| Video sample filtering that can be used to train the ML model | Video-signal processing | Not Started |
| Develop and test ML Model to take filtered video samples and corresponding labels as input and predict the non-verbal skills and areas of improvement | Machine Learning and Deep Learning Algorithms | Not Started |
| Conceptualize and integrate the designed ML model with an Android app (Deployment) | Android Development | Not Started |

**Details:**

While the key to success in both personal and professional relationships lies in our ability to communicate well, it’s not just the words that we use but our nonverbal cues or “body language” that speak the loudest. Our facial expressions, gestures, posture, tone of voice, and level of eye contact are powerful communication tools.

Non-verbal cues frequently reveal the intention of the children and reflect his/her emotional reactions. By improving how children understand and use nonverbal communication, we can help them express what they really mean, help them connect better with others, and build stronger, more rewarding relationships.

The Video-based Test of Communication Skills (VTCS) is an innovative, computer-administered test to understand the many different types of nonverbal communication or body language which include:

1. Facial expressions: The human face is extremely expressive, able to convey countless emotions without saying a word.The facial expressions for happiness, sadness, anger, surprise, fear, confusion and disgust are the same across cultures.
2. Body movement and posture: This type of nonverbal communication includes posture, bearing, stance, and the subtle movements children make.
3. Gestures: Children may wave, point, beckon, or use hands when arguing or speaking animatedly, often expressing themselves with gestures without thinking.
4. Eye contact: Eye contact is also important in maintaining the flow of conversation and for gauging the other person’s interest and response.
5. Voice: It’s not just what children say, it’s how they say it. Their timing and pace, how loud they speak, their tone and inflection, and sounds that convey understanding, such as “ahh” and “uh-huh.” Their tone of voice can indicate sarcasm, anger, affection, or confidence.

## **Assess the developmental needs of the individual child**

**Objective:** To analyse child’s comprehension to the video course content

**Concept:** To develop a tool to understand how children are responding/comprehending to the course video content they watch

**Issues addressed:** “No two children are the same”. Some learn quickly; others are slow. How do we assess the developmental needs of the individual children? The Pradigi App provides learning content in the form of videos and games which are same for every children. We need to find ways to make learning "relevant, authentic, and valuable" in a child’s lives. This can be done if we can interpret comprehension level of students to the same video by analysing their non-verbal communication like facial expressions etc.

**Tech Needs:** A list of technology know-how or clear projects.

|  |  |  |
| --- | --- | --- |
| **Projects** | **Know-how** | **Status** |
| Project Conceptualization - which non-verbal skills to look for and what should be the labels | Non-Verbal Communication Expert | Not Started |
| An android app to collect labeled dataset to be used for testing and training the ML model | Android Development | Not Started |
| Video sample filtering that can be used to train the ML model | Video-signal processing | Not Started |
| Develop and test ML Model to take filtered video samples and corresponding labels as input and predict the non-verbal skills and areas of improvement | Machine Learning and Deep Learning Algorithms | Not Started |
| Conceptualize and integrate the designed ML model with an Android app (Deployment) | Android Development | Not Started |

**Details:**

Facial expression is the most frequently used nonverbal communication mode by the children and facial expressions of the students are significantly correlated to their emotions which helps to recognize their comprehension towards the course content. Student’s facial expressions can be used as a valuable source of feedback. It can be used to determine whether or not to slow down, speed up, or in some other way modify the presentation of content. Recognition of emotions from facial expressions involves the task of categorizing active and spontaneous facial expressions to extract information about the underlying emotional states. Momentary expressions that signal emotions include muscle movements such as raising the eyebrows, wrinkling the forehead, rolling the eyes or curling the lip. When students are feeling uncomfortable, they may have lowered brow, drawn together brow, horizontal or vertical forehead wrinkles, and have a hard time in maintaining eye contact.

We need a tool which can evaluate these facial expressions to understand how the child is comprehending to the content they are viewing. This will help us implement TaRL (Teaching at the Right Level) by recommending content according to their comprehension level.

Source: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3795200/>

## **Custom feature extraction from video assignments**

**Concept:** To extract features from the videos made by children in different assignments given to them.

**Objective:** Evaluate the videos made by children on different topics

**Issues addressed:** Instead of only watching and playing, we at Pratham encourage kids to craft their own content and make videos on different topics which we give to them. The problem is then to evaluate all those videos generated by children. We need an app to help us extract features of a video which will give an overall idea about the video and its relevance with the topic given.

**Details:** Kids actually love to express themselves, but sometimes they feel like they don't have much of a voice. Encouraging kids to be more of a maker is just a matter of pointing to someone or something they admire and giving them the technology to make their vision come alive. With the right tools, the child can be using that time to build creative skills while sharing their stories, opinions, and ideas. They can play around with animated storytelling apps that let them record a mini-movie with movable characters, props, and settings. Once, they make the videos they can upload it to a platform to get feedback. Having a real audience shows kids that their efforts can matter, so children can use app to get feedback on how they are doing and how can they improve on. The basic aim of the feature extraction app will be to detect faces, objects, texts (if any) in the video submitted by the child and then come up with a report of how relevant was the video with the topic.

## **Object Identification: Biodiversity App**

**Concept:** Children can check out local biodiversity wherever they are with an app that says all about the species of animals and plants they find.

**Objective:** To help children enhance their interest, awareness and knowledge of their local biodiversity

**Issues addressed:** Children are natural born explorers. Children wonder which species of bird it is that they keep seeing in their village. They want to know more about the mammals that call their surroundings. They keep asking their parents or elders nearby but get limited information about the same. We need an app that can help children quickly and easily identify flora and fauna, record their findings and learn more about them.

**Details:**

With a novel modelling and mapping platform covering tens of thousands of species -- everything from mammals and birds to plants, amphibians, reptiles, and fish, the app presents localised species detailed information. This system will enable children search through and learn more about the species they just spotted. They can just take photo and the app will run wikipedia for them. Moreover, it can be used to scroll through the photographs, learn more about the species, and also see which more species live nearby their villages.

## **Object Identification: Word Pirate App**

**Concept:** To identify text from different objects and save those words as part of building the vocabulary

**Objective:** To identify text from different objects and save those words as part of building the vocabulary

**Issues addressed:** Building a language vocabulary is a daunting task. We need a fun way to keep the kids inquisitive to keep searching for new words and understanding their meaning and context, hence keep building their vocabulary.

**Details:** Every day children come across a lot of objects which have text written over them. If they just start paying attention to these text and their meaning, it can help build a child’s vocabulary. The app intents to generate curiosity in children to look for new texts on different objects they come across, click a picture of the object and upload on the app. The app will identify the text written over it, give its meaning and also translation in the native language. They can then save those words and start building their vocabulary.

# **NLP/Chatbot**

1. Conversation skills (Indic languages)
2. Evaluating long form answers - Training platform

## **Conversation skills in Indic languages**

**Concept:** Develop a tool where children can learn verbal communication skills through exposure to conversation flows

**Objective:** Develop a tool where children can learn verbal communication skills through exposure to conversation flows

**Issues addressed:** Verbal communication skills improve if real time feedback is provided to the learner. The unavailability of teachers/trainers in rural parts of the country is a big barrier to achieving this important job skill.

**Details:**

The app allows users to learn and practice their language skills by learning new words and grammar lessons through conversation. By having a chatbot with which they can have a conversation, children can not only build their vocabulary but also understand how to use the words in a conversation by practising it. The basic aim is to expose learner to conversation scenarios on a speech enabled chatbot like tool. Then use Natural language processing to provide real time feedback on pronunciation, grammar, choice of words etc. for learner response.

## **Computer Assisted Assessment (CAA) of descriptive based examination**

**Concept:** To develop an automated system which can evaluate essays or long form answers in Indic languages

**Objective:** To develop an automated system which can evaluate essays or long form answers in Indic languages

**Issues addressed:** Essay questions designed to measure writing ability, along with open-ended questions requiring short/long answers, are highly-valued components of effective assessment programs, but the expense and logistics of scoring them reliably often present a barrier to their use.

**Details:**

Automatic evaluation is preferred to manual assessment to avoid monotonic, bias errors and conserves time for other activities. Hence automatic assessment is vital for educational system. Keyword analysis, natural-language processing and Information mining techniques are the main approaches adopted for text assessment.

# 

# **UX enhancement**

1. AR